DISCUSSION PAPER SERIES

IZA DP No. 9714

Smoking Peer Effects among Adolescents: Are Popular Teens More Influential?

Juan David Robalino

February 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Smoking Peer Effects among Adolescents: Are Popular Teens More Influential?

Juan David Robalino
Cornell University
and IZA

Discussion Paper No. 9714
February 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9714
February 2016

ABSTRACT

Smoking Peer Effects among Adolescents: Are Popular Teens More Influential?

In this paper I analyze adolescent peer effects on cigarette consumption while considering the 'popularity' of peers. The analysis is based on AddHealth data, a four wave panel survey representative of American high-school students. The data include the social network of each school, which we use to measure peers' popularity from network centrality measures, in particular weighted-eigenvector centrality. We use lagged peers' behavior at the grade level to alleviate potential homophilic confounds, and we include school fixed effects to control for contextual confounds. We find that most of the aggregate peer effects regarding cigarette smoking come from the smoking propensity of the 20% most popular kids, suggesting a mediation from social status. This effect persists seven and thirteen years later (wave 3 and 4 of the data). Indeed, the smoking propensity of the bottom 80% seems to have a negative influence on the probability of smoking in the long run (wave 3 and 4). These results hint to the importance of knowing not only the smoking propensity within a school but also the place of the smokers within the social hierarchy of the school.

JEL Classification: I1

Keywords: peer effects, status

Corresponding author:

Juan David Robalino
Department of Economics
Cornell University
Ithaca, NY 14850
USA
E-mail: jr872@cornell.edu

1 Introduction

More than 480,000 deaths are attributed to cigarette smoking every year in the US alone (Center for Disease Control and Prevention). There is no question that the consumption of cigarettes is an important welfare concern and that understanding the patterns driving smoking incentives is of great importance. In particular, initiating smoking at a young age is correlated with smoking more cigarettes per day and with a lower probability of quitting later in life, emphasizing the importance of understanding the drivers of smoking behavior among youth (Everett et al., 1999; Lando et al., 1999). Furthermore, smoking seems to have an important social component, specially for adolescents who can be particularly vulnerable to influence as they try to fit in with their peers (Aral and Walker, 2012; Gardner and Steinberg, 2005). Indeed, a large number of studies find peer effects among adolescents regarding cigarette smoking (e.g., Clark and Lohéac, 2007; Gaviria and Raphael, 2001; Krauth, 2006; Lundborg, 2006; Nakajima, 2007; Powell et al., 2005; Valente et al., 2005). These effects may have many mechanisms, such as social pressure or an easier supply of cigarettes through peers. Another mechanism may be that the behavior of peers may signal the social status associated with smoking. If so, the 'popular' teens may be most influential. Thus, this paper aims to distinguish the influence from peers with distinct levels of popularity.

The present study is based on data from the National Longitudinal Survey of Adolescents Health (AddHealth), a four-wave panel (1995, 1996, 2002 and 2008) capturing a representative sample of American High Schools. This data set is ideal as it contains detailed information about substance use as well as rich social network data, which we use to define the popularity of students by means of network centrality measures. We will alleviate homophily and contextual confounds by using the whole grade as peer referent, peers' lagged behavior, as well as school fix effects.

We find that higher mean popularity of peer smokers (i.e., higher network centrality measures) increases the probability of an individual picking up smoking; analogously, higher mean popularity of non-smokers decreases the probability of an individual smoking later on. Perhaps more interesting is that, in contrast with aggregate peer effects as previously reported in the literature, these patterns persists seven and thirteen years after peers' behavior was measured (i.e., in wave 3 and 4 of the data). Similarly, by decomposing the smoking propensity of peers into the propensity of the 20% most popular and that of the 80% least popular, we find that peer effects are mainly driven by the 20% most popular teens. Furthermore, in the long run we find a negative influence from the smoking propensity of 80% least popular peers (in wave 3 and 4 of the data). Similar results apply to the number of cigarettes smoked per month in 2008 (wave 4), as well as the age of initiation. These patterns suggest that influence may be mediated by the social status of peers: individuals seem to imitate the smoking behavior of popular peers and avoid the behavior of unpopular peers.

The remainder of this paper is organized as follows: In Section 2 we present an overview of the literature. Section 3 discusses the data and our estimation strategy. Section 4 presents our results and Section 5 concludes.

2 Overview

2.1 Difficulties with peer effects estimation

Manski (1993) uncovers the many difficulties of peer effects estimation, which he referres to as *the reflexion problem*. One important issue is that a correlation between the group's average behavior and an individual's behavior can be attributed to three mechanisms:

- *endogenous effects* where an agent's choice is influenced by the choices of the group
- *exogenous (contextual) effects* where individuals in a given group may behave similarly

because the whole group experienced an (unobserved) exogenous shock

- *correlated effects* where individuals in a group behave similarly because they have similar unobservable characteristics and self-select into the group.

Peer effects are concerned with the endogenous mechanism, but often observational data make it difficult to disentangle endogenous effects from contextual and correlated effects.

Several approaches have been used to attempt a clean identification. Sacerdote (2001) set the standard in the economics literature in a study regarding peer effects on the Grade Point Average (GPA) of college students. His strategy consists of using exogenous group formation from random allocation to college dorms in order to control for peers' self-selection and homophily. He also uses peers' lagged GPA (from high-school) to control for contextual confounds.

A more criticized approach is that of Christakis and Fowler (2007): using data from the The Framingham Heart Study consisting of a large (sparse) social network, the authors rely on the direction of social ties nominations and the panel aspect of the data (the behavior of nominated peers at time t to predict the behavior of the ego at time $t+1$) to capture peers influence on the diffusion of obesity and subsequently on the diffusion of smoking Christakis and Fowler (2007, 2008). Yet, their approach has been subject to a lot of criticism by Cohen-Cole and Fletcher (2008) among others, who argue that Christakis and Fowler (2007) do not properly account for contextual effects and homophily.

More sophisticated approaches use Monte Carlo estimation to identify the influence from peers. Two main types of models exist: The first are structural models from game theory where the influence of peers is estimated as the equilibrium outcome of strategic interactions (Krauth, 2006; Nakajima, 2007); the second type consist of agent-base models capturing the co-evolution of friendship formation and peers' influence (Mercken et al., 2010; Steglich and Snijders, 2010).

2.2 Our place in the literature

We will follow on the work from economists on smoking peer effects using the lagged behavior of peers at the grade level to alleviate homophily and including school fix effects to control for contextual confounds.

For example, looking at smoking peer effects Lundborg (2006) uses classmates as peer referents, Clark and Lohéac (2007) use school grade, Gaviria and Raphael (2001) and Powell et al. (2005) use the whole school, and Norton et al. (1998) use neighborhoods. Regarding contextual cofounds, Gaviria and Raphael (2001) and Powell et al. (2005) control for cigarette prices and public policy variables, while Clark and Lohéac (2007) and Lundborg (2006) include school fix effects. All these papers find strong evidence of peers influence on cigarettes use. Gaviria and Raphael (2001), Powell et al. (2005) and Norton et al. (1998) also instrument the behavior of peers using parents behavior (e.g., smoking of peers' parents should affect peers' smoking behavior but not the smoking behavior of other individuals). They find that IV estimations affect very little the peer effect estimates, hence we do not use IV estimations in our analysis. It is worth mentioning that studies using Monte Carlo estimation also find strong evidence of peer effects Mercken et al. (2010); Nakajima (2007).[1] Since we will not follow their approach, we will not discuss them further.

Several studies have analyzed the link between popularity of adolescents and their smoking behavior. For example, Alexander et al. (2001) find that in schools with smoking rates above average, popular students are more likely to smoke; while in schools with smoking rates below average, the opposite is true. Similarly Michell and Amos (1997) use qualitative data from Scottish school girls and find that popular girls smoke to maintain their image while unpopular girls smoke with the hope of gaining social status, but found no effect in the mid-popularity range.

[1] Krauth (2006) uses Monte Carlo estimation and found that other studies overestimate peer effects. Yet, he used peers' behavior reported by the ego, and Norton et al. (2003) finds that such measures result in inconsistent estimations when compared with estimations with actual peers' behavior.

Valente et al. (2005) and Ennett et al. (2006) find a positive relationship between popularity (measured by the number of incoming friendship nominations) and smoking. These papers hint at the role of social status and how smoking may be use as a strategy to climb the social ladder.

We follow the reverse logic by testing how the popularity of peers affects their influence on individuals. We posit that the behavior of popular teens can be associated with social status as opposed to them choosing their behavior to gain status. In other words, we posit that the most popular teens are popular regardless of their smoking behavior, yet, their choices will define whether smoking will be associated with social status for the rest of the students. Thus, individuals may be driven to imitate popular peers in an implicit quest for social status. To the best of my knowledge, this is the first paper analyzing the role of peers' popularity and smoking behavior on their influence towards individuals' smoking choices.[2]

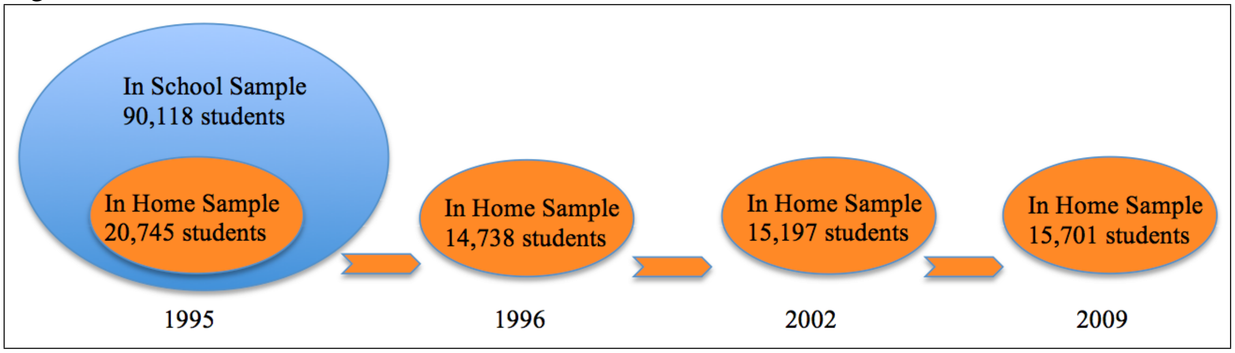
3 Data and Identification Strategy

3.1 Data

We used data from the National Longitudinal Survey of Adolescents Health (AddHealth), a representative sample of high schools in America. This data set contains detailed information about adolescents' substance use as well as household and parents' characteristics and detailed information on the social network of friends. The In-School survey conducted in 1995 covered 90,118 adolescents in 144 schools. A representative sub-sample of 20,745 students also completed the 1995 In-Home wave 1 survey which included more detailed information such as family income and parents' smoking habits. Finally, about 15,000 students were surveyed again in 1996 (In-Home survey wave 2), as well as in 2002 (In-Home wave 3) and in 2008 (In-Home wave 4; see Figure 1).

[2] In his book "The Tipping Point", Gladwell (2006) presents informal interviews from smokers describing their first impressions about cigarettes; most descriptions point towards a given person whom they admired in their youth and who smoked. In fact, this paper is inspired by Gladwell (2006).

Figure 1: Data structure.



Our general strategy is to regress smoking in each of the last three panel waves on peers' smoking in the first panel wave (the 1995 In-School survey) and including school fixed effects. Using the exhaustive In-School sample to control for peers' behavior provides a rich picture of the reference group. After excluding individuals with missing information and schools with insufficient social network data, we end up with a sample of about 77,000 peers and a core sample of about 7,500 individuals.

3.2 Popularity measures

To identify the popularity of peers we use the In-School friendship nomination. Each participant was asked to identify their five closest male friends and their five closest female friends in order of importance (i.e., best friend, second best friend, etc). This allows us to construct an adjacency matrix of friendship A_s for each school and to compute the centrality measures for individuals within each school.[3] In particular, we consider in-degree, out-degree, eigenvector centrality and Katz–Bonacich centrality (described in detail below; see Jackson (2008) Chapter 2 for a discussion about network centrality measures).[4]

[3]Matrix A_s has entries $a_{ij} = 1$ if individual i nominated individual j as a friend and $a_{ij} = 0$ otherwise; $a_{ii} = 0$ for all i.

[4]Note that about 15% of students were not properly matched to the nomination roster and thus their incoming nominations were lost; we consider these peers in our robustness checks.

Degree centrality In- and out-degree centralities measure the number of incoming and outgoing friendship nominations respectively. In-degree is the most straightforward measure of popularity, capturing simply how many people nominated the given individual as a friend. We are less interested in the out-degree centrality since any student may claim a friendship with any other.

Eigenvector centrality This centrality measure is of particular interest as it captures the importance of an individual. The intuition is that the eigenvector centrality of individual i is proportional to the centrality of peers who nominated i as a friend, hence being nominated by central people makes you more central. In terms of adolescents, we define individual i as popular when the other popular teens consider him/her a friend. More formally, the centrality v_i of individual i in school s is

$$v_i = \frac{1}{\lambda} \sum_j v_j a_{ji}$$

or equivalently using vector notation we recover the eigenvector expression $\lambda v = A'v$, where λ is the largest eigenvalue in order to ensure that all the elements of v are positive (as guaranteed by the Perron–Frobenius theorem).[5]

To account for the difference between being the best friend of many people as opposed to being the fifth best friend of many people, we also compute a weighted-eigenvector centrality by replacing the non-zero entries of adjacency matrix by $a_{ij} = 1/n_{ij}$ where n_{ij} is the order of nomination (i, j) among all the nominations of individual i. The weighted-eigenvector centrality is our preferred popularity measure.

Katz–Bonacich There are several alternative measures of centrality designed to capture status or power of individuals. One widely used measure is the Katz–Bonacich centrality: individuals

[5] Note that we use the transpose of A so that the centrality is proportional to the centrality of the peers who nominate individual i, hence, ignoring individuals claiming to be friends with the popular teens.

have a base centrality proportional to their in-degree, $\alpha A'\mathbf{1}$, and part of the centrality of their friends and their friends' friends, etc., gets transferred discounted by parameter β resulting in the following centrality measure:

$$\begin{aligned} c &= \alpha A'\mathbf{1} + \beta A'[\alpha A'\mathbf{1}] + [\beta A']^2[\alpha A'\mathbf{1}] + ... \\ &= [I - \beta A']^{-1}\alpha A'\mathbf{1} \end{aligned}$$

This expression is well defined provided that parameter β is smaller than the inverse of the largest eigenvalue of matrix A' and, thus, we set β equal to this value. For simplicity, we set $\alpha = 1$.

More precisely, we will consider a weighted version of this centrality measure by replacing the non-zero entries of adjacency matrix A with the inverse of the nomination order of friends.

Normalization of centrality measures These measures are sensitive to network size and hence they are not directly comparable across schools. To normalize our measures we take two approaches: the first one is to standardize the popularity measures; the second one is to classify individuals into quintile groups for each grade level based on their centrality in order to identify, for example, the 20% most popular teens in each grade.

3.3 Variables

The questions about smoking are slightly different in the In-Home questionnaire (for the dependent variables) than in the In-School questionnaire (for the influence variables). For individuals in the In-Home survey wave 1, 2 and 3, we consider two main measures of smoking: having had at least one cigarette in the 30 days prior to the interview; and having smoked every day during the month prior to the interview. In wave 4 (recall this interview took place thirteen years after wave 1) we consider the following two measures: having ever tried cigarettes; and having ever smoked every day for at least one month. In wave 4 we also consider the number of cigarettes per month, the age at which the individuals tried their first cigarette and the age at which individuals started smoking every day.[6] The top right panel of Table 1 summarizes these measures. In 1996, 32% of students tried cigarettes and 11% smoked every day; by 2008 these figures were 63% and 43% respectively. Our estimation strategy requires that the correlations between smoking in 1995 and in subsequent years not be too high. In Appendix A1 we present these correlations and the first column shows that there is considerable variation between smoking in 1995 and subsequent years (ranging from a correlation with number of cigarettes per month in 2008 of 0.29, to a correlation with trying cigarettes in 1996 of 0.53).

For peers in the In-School survey, we define as smokers those who reported smoking "once or twice a week" to "daily". We purposely use a high threshold of consumption for peers because we suspect that influence arises from regular users who may influence an (initially) soft consumption to the newly initiated individuals which may eventually result in regular smoking later on. The bottom right panel of Table 1 summarizes the smoking patterns of peers classified by their standardized weighted-eigenvector centrality. In particular we note that, on average, there is no difference in the smoking propensity of the 20% most popular and the 80% least popular (both have a smoking propensity of 0.16). Also note that the normalized popularity of smokers and non-smokers is essentially zero implying that both of these groups have average popularity. We also report the correlations between our popularity measures in Appendix A2, which capture the difference between the distinct popularity measures.

Finally, the top left panel of Table 1 summarizes the demographic covariates used in our analyses. By design, this sample is representative of America high schools in 1995.

[6]Number of cigarettes is compute by multiplying the answers to the questions "During the past 30 days, on how many days did you smoke cigarettes?" and "During the past 30 days, on the days you smoked, how many cigarettes did you smoke each day?"

Table 1: Summary Statistics.

	Mean	SD	Min	Max	N		Mean	SD	Min	Max	N
Demographics						***Smoking***					
Male	0.49	0.50	0	1	7620	Tried by 1995	0.25	0.43	0	1	7620
Age	16.32	1.61	12.08	21	7620	Tried 1996	0.32	0.47	0	1	7620
White	0.58	0.24	0	1	7620	Every day 1996	0.11	0.31	0	1	7620
Black	0.20	0.40	0	1	7620	Tried 2002	0.32	0.47	0	1	6130
Hispanic	0.15	0.36	0	1	7620	Every day 2002	0.17	0.37	0	1	7620
Asian	0.05	0.22	0	1	7620	Tried by 2008	0.63	0.48	0	1	6279
Other	0.01	0.11	0	1	7620	Everyday by 2008	0.43	0.49	0	1	6290
Foreign	0.07	0.25	0	1	7620	# cig./month 2008	95.69	209.46	0	3000	6244
HH income ($1000)	46.53	52.57	0	999	7620	Age first cig.	15.84	3.49	5	31	3955
Move partly for school quality	0.49	0.50	0	1	7620	Age smoked every day	17.38	3.38	5	30	2661
Mother Smokes	0.29	0.45	0	1	7620						
Father Smokes	0.25	0.43	0	1	7620	***Peers***					
Cigs at home	0.31	0.46	0	1	7620	Smoking propensity of					
Out of Sch. in 1996	0.05	0.22	0	1	7620	whole grade	0.16	0.09	0.01	0.59	7620
New student in 1995	0.61	0.49	0	1	7620	20% most pop.	0.16	0.14	0	1	7620
Weekly earnings	53.14	83.83	0	900	7620	80% least pop.	0.16	0.09	0	0.61	7620
8th Grader	0.15	0.36	0	1	7620	unmatchable peers	0.18	0.15	0	1	7551
9th Grader	0.22	0.41	0	1	7620	Mean pop. of smokers	0.04	0.76	-0.29	9.22	7620
10th Grader	0.23	0.42	0	1	7620	Mean pop. of non-smokers	0.09	0.49	-0.29	3.28	7620
11th Grader	0.21	0.41	0	1	7620						
12th Grader	0.04	0.20	0	1	7620						

Peer smokers are those who smoke at least “once or twice a week”. Popularity is defined by standardized weighted-eigenvector centrality.

3.4 Econometric Specification

Most teens go to their local public school and many parents have little flexibility to switch neighborhoods based only on school choices. Even when they consider the schools in the neighborhood, they will probably consider many other aspects of the school like academic rank and facilities before considering the smoking propensity of the school. Additionally, our variation is within schools; peers' behavior at the schools' grade level should be at least partially exogenous and should partially correct for homophily. We also control for new students and for parents who claimed to have chosen their neighborhood in part for the schools quality. Furthermore, we use lagged peers' behavior which will help us deduce the direction of influence and further alleviate homophily confounds. Finally, school fixed effects will control for contextual factors such as price of cigarettes as well as school's implementation of smoking restrictions and the general local sentiment towards smoking.

Our main econometric specification will be a logistic regression of the smoking behavior $Y^t_{i,s,g}$ of individual i at time t on the mean popularity of smokers, $\bar{P}^{t_0}_{s,g,y=1}$, and the mean popularity of non-smokers, $\bar{P}^{t_0}_{s,g,y=0}$, controlling for the smoking propensity, $\bar{Y}^{t_0}_{s,g}$, at the grade level g in school s during the first wave, t_0. Formally we estimate the following equation:

$$P(Y^t_{i,s,g} = 1 | X_i, s, g) = \frac{e^{\mu_{i,s,g}}}{e^{\mu_{i,s,g}} + 1}$$

$$\mu_{i,s,g} = c + z_s + \beta X_i + \alpha_1 \bar{P}^{t_0}_{s,g,y=0} + \alpha_2 \bar{P}^{t_0}_{s,g,y=1} + \alpha_3 \bar{Y}^{t_0}_{s,g} \tag{1}$$

where X_i is a vector of demographics and household characteristics and z_s is a school fixed effect. Our hypothesis is that $\alpha_2 > 0 > \alpha_1$. That is, the more popular peer smokers are, the more likely an individual is to smoke in the future; and the more popular peer non-smokers are, the less likely an individual is to smoke in the future.

We also use two alternative specifications. In the first one we consider the smoking propensity of the 20% most popular teens, $\bar{Y}^{t_0}_{s,g,Pop}$, and the smoking propensity of the 80% least popular teens, $\bar{Y}^{t_0}_{s,g,noPop}$, in school s and grade g:

$$logit(Y^t_{i,s,g}) = c + z_s + \beta X_i + \alpha_1 \bar{Y}^{t_0}_{s,g,Pop} + \alpha_2 \bar{Y}^{t_0}_{s,g,noPop} \tag{2}$$

Here our hypothesis becomes $\alpha_1 > \alpha_2$ capturing a stronger influence from the popular teens.

In the last specification we consider the smoking propensity $\bar{Y}^{t_0}_{s,g,q}$ of the teens in each popularity quintile q in school s and grade g:

$$logit(Y^t_{i,s,g}) = c + z_s + X_i \beta + \sum_{q=1}^{q=5} \alpha_q \bar{Y}^{t_0}_{s,g,q} \tag{3}$$

and we hypothesize that $\alpha_5 > \alpha_{q \neq 5}$. These last two specifications will also allow us to analyze potential negative influence from the least popular teens (i.e., individuals may do the opposite from them).

4 Results

4.1 Benchmark results

Aggregate peer effects We begin by replicating the aggregate peer effects previously found in the literature. Table 2 presents the logistic regression of the probability of trying cigarettes and smoking regularly in 1996, 2002 and 2008. The first row shows the aggregate smoking peer effect from the smoking propensity in the grade. We find strong aggregate peer effects in 1996, where an increase of 10 percentage points in the propensity of smoking in the grade is comparable to the effect of the individual's father smoking. Yet, this peer effect vanishes by 2002 and 2008.

The effects of the remaining covariates are as expected. Having already tried cigarettes in 1995 considerably increases the probability of smoking later on. Males are not more likely

Table 2: Logistic regression of probability of smoking.

	1996		2002		2008	
Smoked	Tried	Every day	Tried	Every day	Ever Tried	Ever Every day
% Regular	1.380**	2.453***	-0.508	-0.499	-0.34	-1.004
smokers in grade	(0.647)	(0.929)	(0.676)	(0.702)	(0.700)	(0.661)
Had tried	2.424***	2.818***	1.632***	1.230***	2.364***	2.044***
smoking in 1995	(0.066)	(0.106)	(0.070)	(0.070)	(0.105)	(0.075)
Male	-0.013	0.004	0.416***	0.279***	0.493***	0.376***
	(0.060)	(0.094)	(0.062)	(0.067)	(0.060)	(0.059)
Age	1.257***	1.776***	0.349	0.675*	1.319***	1.188***
	(0.364)	(0.645)	(0.351)	(0.401)	(0.348)	(0.352)
Age sq.	-0.038***	-0.052***	-0.014	-0.024**	-0.043***	-0.039***
	(0.011)	(0.019)	(0.011)	(0.012)	(0.011)	(0.011)
White	*Ommitted.*					
Black	-0.795***	-1.681***	-0.509***	-0.992***	-0.573***	-0.372***
	(0.115)	(0.242)	(0.114)	(0.141)	(0.100)	(0.105)
Hispanic	0.047	-0.297	-0.225*	-0.494***	0.088	-0.094
	(0.120)	(0.205)	(0.121)	(0.147)	(0.117)	(0.117)
Asian	-0.14	-0.345	0.05	-0.182	0.141	0.092
	(0.169)	(0.332)	(0.175)	(0.219)	(0.167)	(0.170)
Other	0.129	0.492	-0.275	-0.716**	-0.255	-0.012
	(0.265)	(0.414)	(0.313)	(0.346)	(0.266)	(0.265)
Foreign	-0.15	-0.395	-0.308*	-0.037	-0.240*	-0.192
	(0.143)	(0.334)	(0.169)	(0.200)	(0.137)	(0.157)
Out of school	0.398***	0.914***	0.314**	0.562***	0.313**	0.451***
in 1996	(0.136)	(0.167)	(0.144)	(0.144)	(0.159)	(0.150)
New student	0.046	0.013	0.017	-0.098	-0.007	0.123
in 1995	(0.074)	(0.113)	(0.077)	(0.083)	(0.076)	(0.075)
Weekly earnings	0.068*	0.177***	0.044	0.009	0.013	0.087**
($100)	(0.037)	(0.055)	(0.039)	(0.043)	(0.042)	(0.040)
Household Income	-0.552	-4.356*	-0.464	-2.668**	0.602	-1.794**
($1000 000)	(0.631)	(2.572)	(0.711)	(1.313)	(0.710)	(0.785)
Moved partly for	-0.082	-0.187*	-0.017	-0.104	-0.006	-0.073
school quality	(0.063)	(0.096)	(0.065)	(0.071)	(0.063)	(0.062)
Mother smokes	0.036	0.366***	0.175**	0.113	0.049	0.188**
	(0.077)	(0.111)	(0.078)	(0.082)	(0.079)	(0.076)
Father smokes	0.176**	0.245**	0.173**	0.253***	0.034	0.049
	(0.073)	(0.103)	(0.074)	(0.079)	(0.075)	(0.073)
Cigarettes at home	0.298***	0.651***	0.262***	0.280***	0.220***	0.354***
	(0.078)	(0.111)	(0.078)	(0.084)	(0.079)	(0.077)
Const.	-12.610***	-19.403***	-3.977	-6.792**	-9.322***	-9.478***
	(3.040)	(5.390)	(2.908)	(3.309)	(2.874)	(2.893)
Chi sq.	1859.631	1176.187	1004.078	883.82	1078.239	1204.775
Degrees of freedom	138	122	139	138	141	140
Pseudo R sq.	0.254	0.366	0.154	0.148	0.175	0.189
N	7803	7264	6292	7770	6450	6458

Regressions include school dummies. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995.

to smoke than females in 1996, but subsequently they become more likely to smoke in 2002 and in 2008. The probability of smoking increases with age at a diminishing rate. Black students smoke less than white students. Those out of the given school after 1995 are more likely to smoke. Smoking likelihood increases with students' weekly earnings and decreases with household income, albeit the latter effect is very weak. Parents smoking and the availability of cigarettes at home increase the likelihood of a student smoking.

Popularity of smokers/non-smokers Our main results are summarized in Table 3 where we present estimates from equation (1). As previously mentioned, in our preferred specification we measure popularity with weighted-eigenvector centrality (the two columns in bold font in Table 3). We find that the mean popularity of smokers increases the likelihood of an individual smoking, and the mean popularity of non-smokers decreases the probability of an individual picking up smoking, while the effect of the aggregate smoking propensity is similar to the results in Table 2. Further more, in contrast to the aggregate peer effect, the effect of peers' popularity persists in 2002 and in 2008 suggesting that in the long run it is more important who were the smokers than how many peers smoked. The top left panel suggest that an increase of a standard deviation in the mean popularity of smokers results in an increase of 0.313 log-odds of trying cigarettes in 1996; and the bottom right panel suggest that the same increase will result in an increase of 0.173 log-odds of smoking regularly in 2008, thirteen years after having interacted with those peers. These effects are quite strong, similar in magnitude to the effects of parents smoking (see Table 2). Patterns are similar when we use alternative popularity measures.

Note that the one smoking measure for which our popularity effects are not statistically significance is smoking every day in 1996 (top right panel), while in these regressions the aggregate peer effect is particularly strong. Note that in 1996 the average age was 17 years old, thus, most smokers could not buy cigarettes legally yet. This may explain why the aggregate smoking propensity was more important for regular smokers in this period as it probably facilitated the supply of cigarettes among peers.

Robustness One possible explanation for our results is that popular peers may not be more influential due to their social status, but simply they reach more people and they have a simple influence, just like other peers, but towards more friends. To test this, we include the smoking propensity of all nominated friends for each individual. Appendix A3 summarizes the results. The smoking propensity among direct friends absorbs the effect of the smoking propensity in the grade and it is still significant in 2002 and in 2008. Yet, the patterns regarding the mean popularity of smokers and non-smokers remain essentially unchanged.

Another issue is that about 15% of peers could not be matched to the friendship nomination and, thus, they do not have centrality measures. We ran regressions including the smoking propensity of this group and results are essentially unchanged.[7]

4.2 Alternative specifications

Smoking propensity of popular peers In Appendix A4 we estimate equation (2). Again, when we use our weighted-eigenvector centrality we find a similar patter: the aggregate peer effect found in Table 2 seems to be driven mainly by the 20% most popular peers.[8] Note the relative magnitude of the effect: If all of the %20 most popular teens in the grade smoked, the log-odds of smoking the following year would increase by 1.227 (statistically significant at the 1% level), while if all of non-popular teens smoked, resulting in a much larger population of smokers, the increase in the log-odds of smoking would not be statistically different from zero. In fact, in 2002 and 2008 the bottom 80% seem to have a negative influence, that is, the more

[7] Results available upon request.

[8] We find a similar pattern when considering the 10% most popular and the 90% least popular peers, albeit statistical significance is some what weaker. Results available upon request.

Table 3: Logistic regression of probability of smoking on popularity of smokers/non-smokers.

Dependent Variable:	Smoked at least once during past month					Smoked every day past month				
	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
1996										
% Smokers in Grade	1.057	1.114*	1.294*	**1.166***	1.075	2.230**	2.184**	2.235**	**2.233****	2.243**
	(0.672)	(0.671)	(0.674)	**(0.673)**	(0.673)	(0.960)	(0.963)	(0.967)	**(0.958)**	(0.961)
Smokers mean pop.	0.352***	0.326***	0.303***	**0.313*****	0.380***	0.191	0.169	0.083	**0.142**	0.163
	(0.101)	(0.117)	(0.072)	**(0.077)**	(0.100)	(0.159)	(0.195)	(0.114)	**(0.108)**	(0.161)
Non-smokers mean pop	-0.244	-0.112	-0.422***	**-0.392*****	-0.374*	-0.085	-0.407	-0.066	**-0.146**	-0.142
	(0.210)	(0.217)	(0.100)	**(0.103)**	(0.191)	(0.304)	(0.332)	(0.144)	**(0.152)**	(0.300)
2003										
% Smokers in Grade	-0.369	-0.336	-0.222	**-0.271**	-0.384	-0.525	-0.493	-0.455	**-0.434**	-0.517
	(0.699)	(0.699)	(0.702)	**(0.698)**	(0.699)	(0.724)	(0.721)	(0.737)	**(0.723)**	(0.723)
Smokers mean pop.	0.260***	-0.094	0.179***	**0.154****	0.249***	0.235**	0.06	0.245***	**0.221*****	0.238**
	(0.094)	(0.116)	(0.066)	**(0.064)**	(0.091)	(0.104)	(0.133)	(0.084)	**(0.080)**	(0.101)
Non-smokers mean pop	-0.488**	-0.02	-0.247**	**-0.267****	-0.497**	-0.251	-0.026	-0.197*	**-0.276****	-0.394*
	(0.220)	(0.227)	(0.102)	**(0.108)**	(0.201)	(0.235)	(0.243)	(0.112)	**(0.116)**	(0.219)
2008										
% Smokers in Grade	-0.345	-0.263	-0.28	**-0.28**	-0.347	-0.835	-0.772	-0.663	**-0.682**	-0.838
	(0.725)	(0.731)	(0.725)	**(0.727)**	(0.726)	(0.689)	(0.685)	(0.688)	**(0.687)**	(0.688)
Smokers mean pop.	0.279***	0.277**	0.180***	**0.222*****	0.255***	0.243***	0.178	0.171***	**0.173*****	0.243***
	(0.100)	(0.117)	(0.069)	**(0.073)**	(0.094)	(0.093)	(0.113)	(0.065)	**(0.065)**	(0.089)
Non-smokers mean pop	-0.334	-0.448*	-0.163	**-0.251****	-0.478**	-0.631***	-0.253	-0.243**	**-0.325*****	-0.613***
	(0.228)	(0.229)	(0.103)	**(0.109)**	(0.204)	(0.214)	(0.219)	(0.100)	**(0.104)**	(0.194)

Regressions include school dummies. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least "once or twice a week" in 1995. Includes all covariates from Table 2.

of these peers smoked, the less likely an individual is to pick up smoking later in life. We also find similar results using alternative popularity measures.

These effects have a large magnitude. For example, if 10% of the 20% most popular teens smoke (that is a 2% of all the students), the effect has a similar magnitude than if a parent smokes.

Smoking propensity by popularity quintiles In Appendix A5 we estimate equation (3). Looking at our preferred popularity measure, we find similar patterns as before, albeit the more granular cut of peer groups makes these patterns less sharp. The top popularity quintile is the main driver of influence in every period, while in 2002 we find a negative influence from the lowest popularity quintile. Looking at results in 2008, it seems that individuals differentiate the most from the second lowest popularity quintile. This may be because the lowest popularity quintile may have been partly disconnected from the social scene and thus forgotten, while the second lowest popularity quintile may have been more connected to the social scene while retaining the lowest social status among the connected students.

Results in these two alternative specification are also robust to the inclusion of the smoking propensity of direct friends, and the smoking propensity of peers with missing incoming friendship nomination data (i.e., peers with no centrality measures).[9]

4.3 Number of cigarettes per month and age of initiation

In Table 4 we regress the number of cigarettes consumed per month in 2008, as well as the age at which individuals tried their first cigarette and the age at which they smoked every day for the first time. To account for non-smokers, we estimate a Tobit censured equation analogous to equation (1). Looking at our preferred popularity measure (column 4) we find that the same patterns reported above regarding peers popularity apply to the number of cigarettes and the age

[9]Results available upon request.

of initiation. Note that the coefficients reported in Table 4 are the marginal effects on the latent variable of the Tobit model. Computing the marginal effects conditional on smoking, we find that a standard deviation above the mean in the popularity of smokers increases consumption by 8.7 cigarettes per month; and the same variation in the popularity of non-smokers decreases consumption by 20.8 cigarettes per month.

Regarding the age of initiation conditional on ever trying cigarettes, we find that a standard deviation above the mean in the popularity of smokers advances the age at which individuals try their first cigarette by 4.8 years; and the same variation in the popularity of non-smokers delays the age at which individuals try their first cigarette by 4.9 years. Similarly, conditional on ever smoking every day, we find that a standard deviation above the mean in the popularity of smokers advances the age at which individuals started smoking every day by 3.4 years; and the same variation in the popularity of non-smokers delays the age at which individuals started smoking every day by 6.3 years. Results are very similar with alternative popularity measures.

In Table A6 we estimate Tobit regressions analogous to model (2). We do not find statistically significant results for the number of cigarettes smoked. Yet, we find the expected patterns regarding age of initiation: a 10% increase in the smoking propensity of the 20% most popular teens (a mere 2% of the total students) advances the age of the first cigarette by 0.27 years and the age first started smoking every day by 0.25 years; a 10% increase in the smoking propensity of the 80% least popular teens (a 8% of the total students) delays the age of the first cigarette by 0.36 years and the age first started smoking every day by 0.63 years. We find similar patterns using peers' smoking propensity by popularity quintiles (see Table A7).

These patterns are of particular importance in light of the existing evidence that smoking at a younger age has a strong impact in the number of cigarettes smoked and the probability of quitting smoking latter in life Everett et al. (1999); Lando et al. (1999). Popularity of peer smokers during the teenage years seems to be an important driver of these findings.

Table 4: Tobit regression of number of cigarettes smoked per month and age of initiation.

	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
	Number of cigarettes last month				
% Smokers in grade	-214.215	-213.866	-170.553	**-186.232**	-214.447
	(146.771)	(146.672)	(146.793)	**(145.457)**	(146.809)
Smokers mean pop.	31.914*	5.135	24.347*	**26.173****	29.293*
	(19.037)	(23.812)	(13.088)	**(12.982)**	(17.682)
Non-smokers mean pop	-89.367*	-50.018	-49.702**	**-61.239*****	-91.706**
	(46.096)	(46.797)	(21.895)	**(22.323)**	(41.579)
	Age first cigarette				
% Smokers in grade	2.423	2.246	2.12	**2.072**	2.467
	(3.573)	(3.580)	(3.597)	**(3.574)**	(3.572)
Smokers mean pop.	-1.211**	-1.236**	-0.899***	**-1.122*****	-1.152**
	(0.474)	(0.629)	(0.334)	**(0.331)**	(0.461)
Non-smokers mean pop	1.7	2.201*	0.825	**1.422****	2.615**
	(1.167)	(1.173)	(0.549)	**(0.580)**	(1.065)
	Age first smoked every day				
% Smokers in grade	6.523	6.183	5.408	**5.55**	6.504
	(4.814)	(4.809)	(4.853)	**(4.808)**	(4.811)
Smokers mean pop.	-1.883**	-1.269	-1.362**	**-1.397*****	-1.852***
	(0.742)	(0.860)	(0.532)	**(0.536)**	(0.713)
Non-smokers mean pop	4.098***	1.569	1.754**	**2.434*****	4.241***
	(1.554)	(1.607)	(0.757)	**(0.783)**	(1.433)

Regressions include school dummies. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least "once or twice a week" in 1995. Includes all covariates from Table 2.

4.4 Heterogeneity

Finally we explored whether our results vary for different type of students or schools. We modify equation (1) to include interaction term $D_{i,s,g}$ as follows:

$$\begin{aligned} logit(Y_{i,s,g}) &= c + z_s + \beta X_i + \alpha_0 \bar{P}^{t_0}_{s,g,y=0} + \alpha_1 \bar{P}^{t_0}_{s,g,y=1} + \alpha_2 \bar{Y}^{t_0}_{s,g} \\ &+ \gamma_0 D_{i,s,g} + \gamma_1 D_{i,s,g} * \bar{P}^{t_0}_{s,g,y=0} + \gamma_2 D_{i,s,g} * \bar{P}^{t_0}_{s,g,y=1} + \gamma_3 D_{i,s,g} * \bar{Y}^{t_0}_{s,g} \end{aligned}$$

We also used analogous modifications of equation (2) and (3).

We had no particular priors regarding demographics. We interacted the following variables in separate regressions: gender, age, age relative to grade mean, physical attractiveness (rated by interviewer), a dummy for new students, a dummy for foreigners and a dummy for parents claiming to have chosen the neighborhood partly for schools' quality. No systematic patterns were found.[10]

Regarding the individual's own network variables, we suspected that individuals may be more influenced by the peers that are just a little more popular. Similarly, we suspected that individuals with fewer reciprocal nominations (i.e., who claim peers as friends but do not receive the corresponding reciprocal nominations) may be more influenceable as this may be indicative of individuals who try to enter a social circle and who may adapt their behavior to fit in. We ran regressions interacting the popularity of the individual (albeit, this brings back issues of homophily where an individual of a given popularity level may be more related to peers of the same popularity level), as well as regressions interacting a dummy for whether the individual's best friend also nominated him as a best friend, and regressions interacting the overall percentage of reciprocal ties from the individual's nominations. We did not found any systematic patterns.[11] All individuals seem to be mainly influenced by the most popular peers.

[10] Results available upon request.
[11] Results available upon request.

Lastly, we suspected that schools with more network density (i.e., the percentage of friendship ties among the maximum number of ties) may capture social cohesion and, thus, higher levels of influence. Again no systematic patterns were found when we interacted school's network density.[12]

This (lack of) findings may suggest that all individuals are vulnerable to the influence of popular peers when it comes to smoking choices.

5 Conclusion

To the best of my knowledge, this is the first paper looking at the role of peers' popularity regarding smoking peer effects. Using rich data from AddHealth, we constructed popularity measures for peers based on centrality measures in the social network of each school, which we considered together with peers' smoking behavior. In order to alleviate confounds from homophily in our peer effects estimates, we used aggregate peers' behavior at the grade level in the first wave of the survey to predict individuals' smoking behavior in subsequent years. We also included school fixed effect to control for contextual confounds.

We systematically find that the mean popularity of smokers strongly increased the probability of an individual smoking later on. Similarly, the popularity of non-smokers strongly decreased the probability of individuals smoking later on. These effects persist even thirteen years after peers' behavior was measured. Looking at these patterns in an alternative way, we find that most of the aggregate peer effects come from the smoking propensity of the 20% most popular peers. We also find that the smoking propensity of the least popular peers has a negative influence on individuals' smoking in the long run (seven and thirteen years latter). Results are in line with the idea that the behavior of popular teens may be mediated by their social status.

[12]For continuous variables, we regressed separately continuous interactions as well as a dummy signaling values above and bellow the median. Results are available upon request.

But also, that the behavior of unpopular teens may be associated with lower social status and therefore avoided by other students.

There is a general consensus in the literature regarding the importance of peer effects, in particular concerning adolescents' substance use and risky behavior. This paper contributes to the literature by showing the importance of the popularity of peers in predicting their influence. Not only it is important to know the smoking propensity in schools but also who are the users in the social hierarchy. For example, our results suggest that two schools where 10% of teens smoke may experience very different evolutions in the future rates of smoking: If those 10% belong to the 20% most popular teens, log odds of smoking for students the next year would increase by 0.6, while if those 10% belong to the bottom 80% of popular teens, the log odds would not have an increase statistically different from zero.

In the future, we plan to perform several extensions of the current work. For instance, it would be interesting to test how the popularity of peers affects influence on other behaviors including other risky behavior such as consumption of alcohol and drugs, stigmatized topics such as sexual activity, as well as positive behaviors such as grades and sports.

References and Notes

Alexander, C., M. Piazza, D. Mekos, and T. Valente (2001). Peers, schools, and adolescent cigarette smoking. *Journal of Adolescent Health 29*(1), 22–30.

Aral, S. and D. Walker (2012, July). Identifying Influential and Susceptible Members of Social Networks. *Science 337*(6092), 337–341.

Christakis, N. A. and J. H. Fowler (2007). The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine 357*(4), 370–379.

Christakis, N. A. and J. H. Fowler (2008). The Collective Dynamics of Smoking in a Large Social Network. *New England Journal of Medicine 358*(21), 2249–2258.

Clark, A. E. and Y. Lohéac (2007, July). “It wasn’t me, it was them!” Social influence in risky behavior by adolescents. *Journal of Health Economics 26*(4), 763–784.

Cohen-Cole, E. and J. M. Fletcher (2008, December). Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis. *BMJ 337*(dec04 2), a2533–a2533.

Ennett, S. T., K. E. Bauman, A. Hussong, R. Faris, V. A. Foshee, L. Cai, and R. H. DuRant (2006). The Peer Context of Adolescent Substance Use: Findings from Social Network Analysis. *Journal of Research on Adolescence 16*(2), 159–186.

Everett, S. A., C. W. Warren, D. Sharp, L. Kann, C. G. Husten, and L. S. Crossett (1999, November). Initiation of cigarette smoking and subsequent smoking behavior among U.S. high school students. *Preventive medicine 29*(5), 327–333.

Gardner, M. and L. Steinberg (2005, July). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental psychology 41*(4), 625–635.

Gaviria, A. and S. Raphael (2001). School-Based Peer Effects and Juvenile Behavior. *Review of Economics and Statistics 83*(2), 257–268.

Gladwell, M. (2006). *The tipping point: How little things can make a big difference*. Little, Brown.

Jackson, M. O. (2008). *Social and economic networks*. Princeton, NJ : Princeton University Press.

Krauth, B. V. (2006, July). Simulation-based estimation of peer effects. *Journal of Econometrics 133*(1), 243–271.

Lando, H. A., D. T. Thai, D. M. Murray, L. A. Robinson, R. W. Jeffery, N. E. Sherwood, and D. J. Hennrikus (1999, December). Age of initiation, smoking patterns, and risk in a population of working adults. *Preventive medicine 29*(6 Pt 1), 590–598.

Lundborg, P. (2006, March). Having the wrong friends? Peer effects in adolescent substance use. *Journal of Health Economics 25*(2), 214–233.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531.

Mercken, L., T. A. B. Snijders, C. Steglich, E. Vartiainen, and H. de Vries (2010, January). Dynamics of adolescent friendship networks and smoking behavior. *Social Networks 32*(1), 72–81.

Michell, L. and A. Amos (1997). Girls, pecking order and smoking. *Social Science & Medicine 44*(12), 1861–1869.

Nakajima, R. (2007). Measuring Peer Effects on Youth Smoking Behaviour. *Review of Economic Studies 74*(3), 897–935.

Norton, E. C., R. C. Lindrooth, and S. T. Ennett (1998). Controlling for the endogeneity of peer substance use on adolescent alcohol and tobacco use. *Health Economics 7*(5), 439–453.

Norton, E. C., R. C. Lindrooth, and S. T. Ennett (2003). How measures of perception from survey data lead to inconsistent regression results: evidence from adolescent and peer substance use. *Health Economics 12*(2), 139–148.

Powell, L. M., J. A. Tauras, and H. Ross (2005, September). The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. *Journal of Health Economics 24*(5), 950–968.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics 116*(2), 681–704.

Steglich, C. and T. Snijders (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological ... 40*(1), 329–393.

Valente, T. W., J. B. Unger, and C. A. Johnson (2005, October). Do popular students smoke? The association between popularity and smoking among middle school students. *Journal of Adolescent Health 37*(4), 323–329.

Appendix

Table A1: Correlation across measures of smoking.

	Tried by 1995	Tried 1996	Every day 1996	Tried 2002	Every day 2002	Tried by 2008	Every day by 2008	#cigs 2008
Tried by 1995	1							
Tried 1996	0.5351	1						
Every day 1996	0.4917	0.5228	1					
Tried 2002	0.3923	0.4567	0.395	1				
Everyday 2002	0.3485	0.3906	0.4266	0.744	1			
Tried by 2008	0.3534	0.4244	0.265	0.4878	0.3767	1		
Everyday by 2008	0.4296	0.4879	0.399	0.6755	0.5595	0.6568	1	
#cigs 2008	0.2903	0.3168	0.3691	0.4822	0.5317	0.3384	0.5095	1

Table A2: Correlation across measures of popularity.

	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
In-degree	1				
Out-degree	0.3688	1			
Eigenvector	0.4825	0.2125	1		
W. Eigenvector	0.442	0.1962	0.8059	1	
W. Katz-Bonacich	0.8629	0.3115	0.5945	0.611	1

Table A3: Logistic regression of probability of smoking on popularity of smokers/non-smokers.

	1996									
Dependent Variable:	Smoked at least once during past month					Smoked every day past month				
Smoking propensity among:	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
% Direct friend smokers	0.791***	0.805***	0.801***	0.797***	0.793***	1.428***	1.432***	1.432***	1.432***	1.429***
	(0.142)	(0.141)	(0.142)	(0.142)	(0.142)	(0.177)	(0.177)	(0.177)	(0.177)	(0.177)
% Smokers in grade	0.561	0.625	0.785	0.67	0.578	1.261	1.196	1.352	1.295	1.271
	(0.699)	(0.698)	(0.701)	(0.701)	(0.700)	(1.015)	(1.013)	(1.027)	(1.015)	(1.015)
Smokers mean pop.	0.383***	0.309**	0.347***	0.359***	0.425***	0.154	0.161	0.101	0.127	0.146
	(0.104)	(0.121)	(0.076)	(0.082)	(0.105)	(0.164)	(0.209)	(0.115)	(0.111)	(0.161)
Non-smokers mean pop	-0.262	-0.044	-0.439***	-0.432***	-0.392*	-0.195	-0.484	-0.146	-0.172	-0.162
	(0.222)	(0.226)	(0.108)	(0.110)	(0.203)	(0.324)	(0.358)	(0.159)	(0.169)	(0.330)
	2002									
% Direct friend smokers	0.912***	0.934***	0.919***	0.918***	0.915***	0.895***	0.902***	0.899***	0.899***	0.898***
	(0.150)	(0.149)	(0.149)	(0.149)	(0.149)	(0.145)	(0.145)	(0.145)	(0.145)	(0.145)
% Smokers in grade	-0.924	-0.941	-0.807	-0.837	-0.94	-1.370*	-1.356*	-1.354*	-1.285*	-1.360*
	(0.733)	(0.733)	(0.737)	(0.733)	(0.734)	(0.770)	(0.769)	(0.784)	(0.770)	(0.769)
Smokers mean pop.	0.230**	-0.153	0.164**	0.138**	0.234**	0.236**	-0.002	0.254***	0.238***	0.257**
	(0.097)	(0.122)	(0.068)	(0.066)	(0.094)	(0.109)	(0.138)	(0.092)	(0.089)	(0.108)
Non-smokers mean pop	-0.399*	0.028	-0.211**	-0.232**	-0.419**	-0.131	0.118	-0.166	-0.264**	-0.276
	(0.228)	(0.234)	(0.106)	(0.112)	(0.209)	(0.241)	(0.253)	(0.118)	(0.123)	(0.226)
	2008									
% Direct friend smokers	0.732***	0.747***	0.744***	0.742***	0.734***	0.940***	0.956***	0.950***	0.949***	0.943***
	(0.177)	(0.177)	(0.177)	(0.177)	(0.177)	(0.152)	(0.152)	(0.152)	(0.152)	(0.152)
% Smokers in grade	-0.866	-0.786	-0.797	-0.799	-0.87	-1.489**	-1.423**	-1.346*	-1.361*	-1.492**
	(0.757)	(0.765)	(0.759)	(0.759)	(0.759)	(0.729)	(0.726)	(0.729)	(0.727)	(0.729)
Smokers mean pop.	0.274***	0.256**	0.171**	0.209***	0.274***	0.263***	0.206*	0.203***	0.200***	0.292***
	(0.101)	(0.121)	(0.070)	(0.073)	(0.097)	(0.098)	(0.117)	(0.069)	(0.069)	(0.095)
Non-smokers mean pop	-0.366	-0.475**	-0.161	-0.226**	-0.510**	-0.654***	-0.3	-0.245**	-0.302***	-0.637***
	(0.239)	(0.239)	(0.108)	(0.114)	(0.215)	(0.227)	(0.228)	(0.104)	(0.109)	(0.206)

Regressions include school dummies. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least “once or twice a week” in 1995. Includes all covariates from Table 2.

Table A4: Logistic regression of probability smoking on smoking propensity by popularity of peers.

Dependent Variable:	Smoked at least once during past month					Smoked every day past month				
Smoking propensity among:	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
1996										
20% most popular	0.692*	0.684*	1.254***	1.227***	1.386***	-0.257	0.5	0.706	0.927**	0.419
	(0.370)	(0.361)	(0.370)	(0.349)	(0.409)	(0.485)	(0.465)	(0.491)	(0.463)	(0.567)
80% least popular	0.589	0.759	0.115	0.125	0.127	2.736***	2.045**	1.702*	1.397	1.994**
	(0.678)	(0.586)	(0.627)	(0.621)	(0.620)	(0.969)	(0.865)	(0.899)	(0.873)	(0.893)
2003										
20% most popular	0.522	0.273	0.926**	0.855**	0.861**	0.654*	0.183	1.350***	1.427***	1.145***
	(0.374)	(0.370)	(0.370)	(0.355)	(0.417)	(0.392)	(0.387)	(0.401)	(0.373)	(0.438)
80% least popular	-0.823	-0.345	-1.126*	-1.076*	-0.949	-1.08	-0.542	-1.635**	-1.812***	-1.285*
	(0.698)	(0.596)	(0.654)	(0.654)	(0.654)	(0.737)	(0.633)	(0.670)	(0.672)	(0.678)
2008										
20% most popular	0.309	0.48	0.58	0.728**	0.899*	0.585	0.103	0.711*	0.787**	1.075**
	(0.413)	(0.388)	(0.377)	(0.370)	(0.464)	(0.398)	(0.360)	(0.367)	(0.352)	(0.439)
80% least popular	-0.378	-0.379	-0.619	-0.849	-0.823	-1.641**	-0.993*	-1.684***	-1.784***	-1.864***
	(0.718)	(0.634)	(0.669)	(0.666)	(0.675)	(0.683)	(0.588)	(0.635)	(0.629)	(0.634)

Regressions include school dummies. Robust standard errors are in parenthesis. . Peer smokers are those who smoke at least “once or twice a week” in 1995. Includes all covariates from Table 2.

Table A5: Logistic regression of probability smoking on smoking propensity by popularity quintiles.

Dependent Variable:	Smoked at least once during past month					Smoked every day past month				
Smoking propensity among:	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
1996										
Top Pop. Quintile	0.727*	0.656*	1.195***	**1.114*****	1.250***	0.05	0.104	0.41	**0.876***	0.278
	(0.400)	(0.395)	(0.379)	**(0.357)**	(0.418)	(0.572)	(0.508)	(0.518)	**(0.472)**	(0.587)
2nd Pop. Quintile	0.973**	1.052**	0.559	**0.468**	0.714*	1.464**	2.108***	0.588	**0.542**	0.764
	(0.466)	(0.445)	(0.393)	**(0.422)**	(0.426)	(0.678)	(0.634)	(0.525)	**(0.577)**	(0.622)
3rd Pop. Quintile	-1.242***	0.356	0.273	**0.453**	-0.011	0.944	0.545	1.823***	**0.938**	1.304**
	(0.452)	(0.404)	(0.395)	**(0.420)**	(0.459)	(0.644)	(0.577)	(0.567)	**(0.606)**	(0.629)
4th Pop. Quintile	0.654*	-0.081	-0.789**	**-0.612**	0.002	-0.034	-0.174	-1.099*	**-0.009**	0.364
	(0.393)	(0.418)	(0.394)	**(0.386)**	(0.420)	(0.551)	(0.540)	(0.581)	**(0.546)**	(0.593)
5th Pop. Quintile	-0.485	-0.229	-0.01	**-0.178**	-0.431	0.064	-0.193	0.292	**-0.068**	-0.108
	(0.426)	(0.428)	(0.330)	**(0.365)**	(0.362)	(0.589)	(0.611)	(0.489)	**(0.510)**	(0.477)
2002										
Top Pop. Quintile	0.398	0.079	0.913**	**0.828****	0.687	0.659	0.151	1.313***	**1.495*****	1.105**
	(0.396)	(0.396)	(0.383)	**(0.365)**	(0.421)	(0.430)	(0.424)	(0.410)	**(0.378)**	(0.440)
2nd Pop. Quintile	0.639	0.1	-0.216	**-0.145**	0.531	-0.054	0.264	-0.729*	**-0.766***	0.021
	(0.491)	(0.464)	(0.385)	**(0.421)**	(0.424)	(0.539)	(0.499)	(0.407)	**(0.425)**	(0.461)
3rd Pop. Quintile	-0.48	-0.136	0.684	**0.374**	-0.217	-0.286	-0.321	0.111	**-0.2**	-0.431
	(0.446)	(0.402)	(0.420)	**(0.425)**	(0.451)	(0.458)	(0.411)	(0.430)	**(0.460)**	(0.468)
4th Pop. Quintile	-0.114	-0.488	-0.448	**-0.328**	-0.699	-0.413	-0.900**	-0.779*	**-0.257**	-0.892*
	(0.403)	(0.415)	(0.408)	**(0.397)**	(0.432)	(0.418)	(0.423)	(0.435)	**(0.416)**	(0.463)
5th Pop. Quintile	-1.159**	-0.013	-0.920***	**-0.920****	-0.459	-0.625	0.124	-0.263	**-0.556**	-0.026
	(0.460)	(0.440)	(0.346)	**(0.387)**	(0.378)	(0.470)	(0.474)	(0.370)	**(0.418)**	(0.388)
2008										
Top Pop. Quintile	0.397	0.514	0.375	**0.654***	0.848*	0.463	0.058	0.659*	**0.757****	0.952**
	(0.425)	(0.410)	(0.387)	**(0.375)**	(0.471)	(0.411)	(0.386)	(0.373)	**(0.358)**	(0.450)
2nd Pop. Quintile	0.835*	0.629	0.122	**-0.066**	0.158	0.638	0.406	-0.06	**-0.413**	0.258
	(0.482)	(0.478)	(0.406)	**(0.424)**	(0.442)	(0.469)	(0.453)	(0.388)	**(0.412)**	(0.412)
3rd Pop. Quintile	-0.696	0.317	0.143	**0.105**	-0.748*	-0.503	-0.048	0.069	**0.348**	-0.464
	(0.473)	(0.419)	(0.457)	**(0.441)**	(0.449)	(0.460)	(0.393)	(0.424)	**(0.417)**	(0.444)
4th Pop. Quintile	-0.517	-0.373	-1.220***	**-1.095*****	0.066	-0.56	-0.262	-1.089***	**-1.329*****	-0.835**
	(0.405)	(0.435)	(0.407)	**(0.395)**	(0.431)	(0.387)	(0.413)	(0.398)	**(0.389)**	(0.423)
5th Pop. Quintile	-0.425	-0.652	0.415	**0.285**	-0.268	-1.239***	-0.766*	-0.502	**-0.269**	-0.701*
	(0.474)	(0.448)	(0.353)	**(0.368)**	(0.394)	(0.436)	(0.430)	(0.347)	**(0.369)**	(0.372)

Regressions include school dummies. Robust standard errors are in parenthesis. . Peer smokers are those who smoke at least “once or twice a week” in 1995. Includes all covariates from Table 2.

Table A6: Tobit regressions on smoking propensity by peers' popularity

	Number of cigarettes last month				
Smoking propensity among	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
20% most popular	-40.717	-101.664	-1.844	3.384	14.568
	(78.769)	(75.465)	(74.875)	(74.171)	(89.907)
80% least popular	-158.269	-132.518	-187.508	-193.401	-200.333
	(143.362)	(124.693)	(130.493)	(128.763)	(131.536)
	Age first cigarette				
20% most popular	-0.496	-1.655	-2.739	-3.895**	-4.203*
	(2.007)	(1.861)	(1.938)	(1.886)	(2.300)
80% least popular	1.192	1.969	3.371	4.875	4.077
	(3.491)	(3.040)	(3.256)	(3.227)	(3.200)
	Age first smoked every day				
20% most popular	-3.669	-0.858	-4.827*	-5.452**	-7.229**
	(2.753)	(2.588)	(2.665)	(2.547)	(3.074)
80% least popular	11.151**	7.015*	11.686***	12.501***	12.710***
	(4.812)	(4.114)	(4.469)	(4.406)	(4.370)

Regressions include school dummies. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least "once or twice a week" in 1995. Includes all covariates from Table 2.

Table A7: Tobit regressions on smoking propensity by peers' popularity quintiles

	Number of cigarettes last month				
Smoking propensity among:	In-degree	Out-degree	Eigen vector	**W. Eigenvector**	W. Katz Bonacich
Top Pop. Quintile	-79.969	-133.667*	-4.542	**7.532**	12.552
	(85.020)	(81.148)	(76.012)	**(74.479)**	(91.277)
2nd Pop. Quintile	76.819	87.199	-118.416	**-99.234**	-39.784
	(108.872)	(92.781)	(81.188)	**(85.564)**	(89.473)
3rd Pop. Quintile	-103.42	-57.084	100.632	**80.162**	-38.868
	(95.040)	(76.659)	(88.673)	**(91.002)**	(96.662)
4th Pop. Quintile	82.019	-92.786	-95.149	**-100.335**	-32.922
	(82.394)	(81.728)	(85.235)	**(79.981)**	(90.604)
5th Pop. Quintile	-189.635**	-130.311	-71.225	**-44.089**	-74.364
	(94.733)	(91.059)	(70.168)	**(74.889)**	(82.212)

	Age first cigarette				
Top Pop. Quintile	-1.207	-2.01	-1.507	**-3.321***	-4.117*
	(2.098)	(2.005)	(1.987)	**(1.906)**	(2.328)
2nd Pop. Quintile	-4.375*	-2.582	-1.298	**0.51**	-0.289
	(2.596)	(2.427)	(2.078)	**(2.162)**	(2.115)
3rd Pop. Quintile	3.037	-1.428	-0.16	**-0.299**	3.47
	(2.479)	(1.905)	(2.279)	**(2.368)**	(2.420)
4th Pop. Quintile	3.205	1.799	6.787***	**5.934*****	-0.559
	(2.035)	(2.143)	(2.097)	**(1.940)**	(2.254)
5th Pop. Quintile	1.731	2.618	-2.257	**-1.706**	1.531
	(2.242)	(2.287)	(1.749)	**(1.873)**	(1.947)

	Age first smoked every day				
Top Pop. Quintile	-3.04	-0.94	-4.421	**-5.064***	-6.580**
	(2.891)	(2.815)	(2.709)	**(2.585)**	(3.122)
2nd Pop. Quintile	-3.798	-1.912	0.048	**2.631**	-1.422
	(3.438)	(3.321)	(2.784)	**(2.990)**	(2.897)
3rd Pop. Quintile	3.3	0.338	-0.412	**-2.36**	3.318
	(3.382)	(2.673)	(3.007)	**(3.136)**	(3.247)
4th Pop. Quintile	4.573*	1.809	7.483***	**8.823*****	5.632*
	(2.697)	(2.929)	(2.888)	**(2.778)**	(3.049)
5th Pop. Quintile	7.886***	4.436	3.526	**2.264**	4.571*
	(3.047)	(3.194)	(2.508)	**(2.719)**	(2.617)

Regressions include school dummies. Robust standard errors are in parenthesis. Peer smokers are those who smoke at least "once or twice a week" in 1995. Includes all covariates from Table 2.